



Figure 2

<i>Question Phrase(s)</i>
"who was"
"how do i"
"where is",
"where can i"
"what are"
"what is"
"what is a"

Figure 3

^ what (is | are | were | does | do | did | should | can) \s
 ^ who (is | are | was | were | did | do | does) \s
 ^ how (to | is | do | did | does | can | would | could | should |) \s
 ^ why (is | do | are | did | were | does) \s
 ^ where (is | was | can | are | were | do | does) \s
 ^ when (is | was | are | were | do | did | does) \s
 ^ which \s

Figure 4

Question Phrase	Candidate Transforms
"what is a"	"the term" "component" "ans" "a computer" "telephone" "collection of" "stands for" "unit"

35

Figure 5

<i>Question Phrase</i>	<i>Candidate Transform</i>	$qt\ f_i$	$w_i^{(1)}$	wtr_i
52 "what is a"	"refers to"	30	2.71	81.3
	"refers"	30	2.67	80.1
	"meets"	12	3.21	38.52
	"driven"	14	2.72	38.08
	"named after"	10	3.63	36.3
	"often used"	12	3.00	36
	"to describe"	13	2.70	35.1

51

Figure 6

<i>Transform Length</i>	<i>Candidate Transform tr_i</i>	wtr_i
3	"is used to"	32.89
	"according to the"	23.49
	"to use a"	21.43
2	"is a"	298.89
	"of a"	94.34
	"refers to"	81.3
1	"usually"	128.23
	"used"	110.39
	"refers"	80.1

65

63

Figure 7

- (1) Examples = RetrieveExamples (QP , numExamples) 71
- for each <Question, Answer> in Examples
- for each candidate transform tr_i 72
- (2) Query = ApplyTransform (Question, tr_i) 72
- (3) Results = SubmitQuery (Query, SE) 73
- for each Document in Results
- docScore = -1
- (4a) Subdocuments = getSubDocuments (Document, subDocLen) 74a
- for each SubDocument in SubDocuments
- (4b) tmpScore = DocumentSimilarity (Answer, SubDocument) 74b
- if (tmpScore > docScore) docScore = tmpScore
- (4c) updateTransformScores (tr_i , docScore) 74c
- updateTransformCounts (tr_i) 74b
- (5) AssignTransformWeights (TransformScores, TransformCounts) 75

Figure 8

Search Engine	Transform	WT_i	Search Engine	Transform	WT_i
AltaVista	"is usually"	377.03	Google	"is usually"	280.68
	"refers to"	373.22		"usually"	275.68
	"usually"	371.55		"called"	256.64
	"refers"	370.14		"sometimes"	253.53
	"is used"	360.07		"is one"	253.24

82

83

Figure 9

